OIPE

RAW SEQUENCE LISTING

DATE: 09/06/2000

PATENT APPLICATION: US/09/645,706

TIME: 12:02:11

Input Set : A:\341-005US1.txt.txt

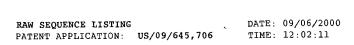
Output Set: N:\CRF3\09062000\I645706.raw

```
4 <110> APPLICANT: Wood, Keith V.
             Gruber, Monika G.
             Zhuang, Yao
      8 <120> TITLE OF INVENTION: SYNTHETIC NUCLEIC ACID MOLECULE
             COMPOSITIONS AND METHODS OF PREPARATION
     11 <130> FILE REFERENCE: 341.005US1
C--> 13 <140> CURRENT APPLICATION NUMBER: US/09/645,706
C--> 13 <141> CURRENT FILING DATE: 2000-08-24
     13 <160> NUMBER OF SEQ ID NOS: 302
     15 <170> SOFTWARE: FastSEQ for Windows Version 4.0
     17 <210> SEQ ID NO: 1
     18 <211> LENGTH: 1629
     19 <212> TYPE: DNA
     20 <213> ORGANISM: Pyrophorus plagiophthalamus
     22 <400> SEQUENCE: 1
     23 atgatgaaga gagagaaaaa tgttatatat ggacccgaac ccctacaccc cttggaagac
     24 ttaacagcag gagaaatgct cttcagggcc cttcgaaaac attctcattt accgcaggct
                                                                             120
     25 ttagtagatg tgtttggtga cgaatcgctt tcctataaag agttttttga agctacatgc
                                                                             180
     26 ctcctagcgc aaagtctcca caattgtgga tacaagatga atgatgtagt gtcgatctgc
                                                                             240
                                                                             300
     27 gccgagaata ataaaagatt ttttattccc attattgcag cttggtatat tggtatgatt
     28 gtagcacctg ttaatgaaag ttacatccca gatgaactct gtaaggtcat gggtatatcg
                                                                             420
     29 aaaccacaaa tagttttttg tacaaagaac attttaaata aggtattgga ggtacagagc
                                                                             480
     30 agaactaatt toataaaaag gatoatoata ottgatactg tagaaaacat acacggttgt
     31 gaaagtette ecaattttat ttetegttat teggatggaa atattgeeaa etteaaacet
                                                                             540
     32 ttacattacg atcctgttga gcaagtggca gctatcttat gttcgtcagg cactactgga
                                                                             600
     33 ttaccgaaag gtgtaatgca aactcaccaa aatatttgtg tccgacttat acatgcttta
                                                                             660
     34 gaccccaggg caggaacgca acttattcct ggtgtgacag tcttagtata tctgcctttt
                                                                             720
     35 ttccatgctt ttgggttctc tataaacttg ggatacttca tggtgggtct tcgtgttatc
                                                                             780
                                                                             840
     36 atgttaagac gatttgatca agaagcattt ctaaaagcta ttcaggatta tgaagttcga
                                                                             900
     37 agtgtaatta acgttccagc aataatattg ttcttatcga aaagtccttt ggttgacaaa
                                                                             960
     1020
     39 qctgaqqttq caqtaaaacg attaaacttg ccaggaattc gctgtggatt tggtttgaca
                                                                            1080
     40 gaatctactt cagctaatat acacagtett ggggatgaat ttaaatcagg atcacttgga
     41 agagttactc ctttaatggc agctaaaata gcagataggg aaactggtaa agcattggga
                                                                            1140
     42 ccaaatcaag ttggtgaatt atgcgttaaa ggtcccatgg tatcgaaagg ttacgtgaac
                                                                            1200
     43 aatgtagaag ctaccaaaga agctattgat gatgatggtt ggcttcactc tggagacttt
                                                                            1260
                                                                            1320
     44 ggatactatg atgaggatga gcatttctat gtggtggacc gttacaagga attgattaaa
                                                                            1380
     45 tataagggct ctcaggtagc acctgcagaa ctagaagaga ttttattgaa aaatccatgt
                                                                            1440
     46 atcagagatg ttgctgtggt tggtattcct gatctagaag ctggagaact gccatctgcg
     47 tttgtggtta aacagcccgg aaaggagatt acagctaaag aagtgtacga ttatcttgcc
                                                                            1500
                                                                            1560
     48 gagagggtct cccatacaaa gtatttgcgt ggaggggttc gattcgttga tagcatacca
                                                                            1620
     49 aggaatgtta caggtaaaat tacaagaaag gaacttctga agcagttgct ggagaagagt
                                                                            1629
     50 tctaaactt
     52 <210> SEQ ID NO: 2
     53 <211> LENGTH: 1626
     54 <212> TYPE: DNA
     55 <213> ORGANISM: Artificial Sequence
```

ENTERED See P.

SEP 20 2000 CIPE/JCWS

9/6/00



Input Set : A:\341-005US1.txt.txt
Output Set: N:\CRF3\09062000\I645706.raw

57 <220> FEATURE:							
58 <223> OTHER INFORMATION: Sequence of clone YG#81-6G01							
60 <400> SEQUENCE: 2							
61 atgatgaago gagagaaaaa tgttatatat ggacccgaac ccctacaccc cttggaagac	60						
62 ttaacagctg gagaaatgct cttccgtgcc cttcgaaaac attctcattt accgcaggct	120						
63 ttagtagatg tggttggcga cgaatcgctt tcctataaag agttttttga agcgacagtc	180						
64 ctcctagcgc aaagtctcca caattgtgga tacaagatga atgatgtagt gtcgatctgc	240						
65 gccgagaata atacaagatt ttttattccc gttattgcag cttggtatat tggtatgatt	300						
66 gtagcacctg ttaatgaaag ttacatccca gatgaactct gtaaggtgat gggtatatcg	360						
67 aaaccacaaa tagtttttac gacaaagaac attttaaata aggtattgga ggtacagagc	420						
68 agaactaatt tcataaaaag gatcatcata cttgatactg tagaaaacat acacggttgt	480						
69 gaaagtette ceaattttat ttetegttat teggatggaa atattgeeaa etteaaaeet	540						
70 ttacatttcg atcetgttga gcaagtggca gctatcttat gttcgtcagg cactactgga	600						
71 ttaccgaaag gtgtaatgca aactcaccaa aatatttgtg tccgacttat acatgcttta	660						
72 gaccccaggg caggaacgca acttattect ggtgtgacag tettagtata tetgeetttt	720						
73 ttccatgctt ttgggttctc tataaccttg ggatacttca tggtgggtct tcgtgttatc	780						
74 atgttcagac gatttgatca agaagcattt ctaaaaagcta ttcaggatta tgaagttcga	840						
75 agtgtaatta acgttccatc agtaatattg ttcttatcga aaagtccttt ggttgacaaa	900						
76 tacgatttat caagtttaag ggaattgtgt tgcggtgcgg	960						
77 gctgaggttg cagcaaaacg attaaacttg ccaggaattc gctgtggatt tggtttgaca	1020						
78 gaatctactt cagctaatat acacagtett agggatgaat ttaaatcagg atcacttgga	1080						
79 agagttactc ctttaatggc agctaaaata gcagataggg aaactggtaa agcattggga	1140						
80 ccaaatcaag ttggtgaatt atgcattaaa ggtcccatgg tatcgaaagg ttacgtgaac	1200						
81 aatqtagaag ctaccaaaga agctattgat gatgatggtt ggcttcactc tggagacttt	1260						
82 ggatactatg atgaggatga gcatttctat gtggtggacc gttacaagga attgattaaa	1320						
83 tataagggct ctcaggtagc acctgcagaa ctagaagaga ttttattgaa aaatccatgt	1380						
84 atcagagatg ttgctgtggt tggtattcct gatctagaag ctggagaact gccatctgcg	1440						
85 tttqtqqtta aacagcccgg aaaggagatt acagctaaag aagtgtacga ttatcttgcc	1500						
86 gagagggtct cccatacaaa gtatttgcgt ggaggggttc gattcgttga tagcatacca	1560						
87 aggaatgtta caggtaaaat tacaagaaag gaacttetga agcagttget ggagaaggeg	1620						
88 ggaggt	1626						
90 <210> SEQ ID NO: 3							
91 <211> LENGTH: 1626							
92 <212> TYPE: DNA							
93 <213> ORGANISM: Artificial Sequence							
95 <220> FEATURE:							
96 <223> OTHER INFORMATION: Sequence of a synthetic luciferase							
98 <400> SEQUENCE: 3							
99 atgatgaaac gcgaaaagaa cgtcatctac ggcccagagc ctctgcaccc attggaagac	60						
100 ctgaccgccg gtgagatgtt gttccgtgct ctgcgtaaac attctcactt gcctcaagcc	120						
101 ctggtggatg tcgtgggcga cgaaagcttg tcttataagg agtttttcga agctactgtc	180						
102 ctqttqqccc aqtctctqca taattqcqqt tacaaaatga acgatgtggt cagcatttgt	240						
103 getgagaata acaccegett tttcatecca gtgattgeeg ettggtacat eggeatgatt	300						
104 gtcgcccctg tgaatgaatc ttatatccca gacgagttgt gcaaggtcat gggtattagc	360						
105 aaacctcaaa tcgtgtttac taccaagaac attctgaata aagtcttgga agtgcagtct	420						
106 cgtactaact teatcaageg cattateatt etggataceg tegagaatat ecaeggetgt	480						
107 gaaagcttgc caaactttat ttctcgttat agcgacggta atatcgctaa cttcaagcct	540						
108 ctgcattttg atccagtgga gcaagtcgcc gctattttgt gctctagcgg cactaccggt	600						
109 etgectaaag gegtgatgea gaeteaceaa aatatetgtg teegettgat teatgeeetg	660						



RAW SEQUENCE LISTING DATE: 09/06/2000 PATENT APPLICATION: US/09/645,706 TIME: 12:02:11

Input Set : A:\341-005US1.txt.txt
Output Set: N:\CRF3\09062000\I645706.raw

110	gacccacgtg	tgggtaccca	gttgatccct	ggcgtgactg	tcctggtgta	cttgccattc	720
111	tttcacgcct	tcggtttttc	tattaccctq	ggctatttca	tggtcggttt	gcgcgtgatc	780
112	atgtttcgtc	gcttcgatca	agaagetttt	ctgaaggcca	ttcaggacta	cgaggtccgt	840
113	agcgtgatca	acqtcccttc	tataatttta	tteetgagea	aatctccatt	ggtcgataag	900
114	tatgacctga	actetttaca	cgaactgtgc	tatagcacta	cccctttggc	taaagaggtg	960
115	gccgaagtcg	ctgccaagcg	tctgaatttg	ccaggtatcc	actacaactt	tggtctgact	1020
116	gagagcacct	ctgctaacat	tcatagettg	cotgatgaat	tcaaatctqq	cagectgggt	1080
117	cgcgtgactc	ctttgatggc	cactaagate	accaaccata	agaccggcaa	agetetgggt	1140
118	ccaaatcaag	tcaacaaatt	atatattaag	ggtcctatgg	tgtctaaagg	ctacqtcaac	1200
110	aatgtggagg	ccactaagga	agctatcgat	gacgatggtt	ggctgcacag	coocgacttt	1260
120	ggttattacg	atgaggagga	acatttctat	atcatagate	gctacaaaga	gttgattaag	1320
121	tataaaggct	ctcaggtcgc	cccagetgag	ctagaagaga	tettaetaaa	gaacccttgc	1380
122	attcgtgacg	tagacatage	aggtatocca	datttddaad	ctagcaaact	gcctagcgcc	1440
122	tttgtcgtga	2202200200	taaggaaatt	accoctaaao	aggtetacga	ctatttggcc	1500
123	gaacgcgtgt	adcaaccayy	atacetacet	accactatec	acttcataga	tagcatccct	1560
	cgcaatgtca						1620
	ggtggc	ccggcaaaat	Lacicylady	gagicgciga	adcagetget	ggaaaaggoc	1626
	<210> SEQ 1	ED NO. 4					2020
	<211> SEQ 1						
	<212> TYPE:		inial Camus				
	<213> ORGAN		iciai seque	iice			
	<220> FEATU		ON Common o	£+1	hatia lugifa	2220	
	<223> OTHER		on: Sequence	e or a synci	decid fucile	erase	
	<400> SEQUE				at at a a a a a a	2++44222424	60
13/	atgatgaaac	gcgaaaagaa	egicatetae	ggeecagage	etetgeacee	actygaagac	120
138	ctgaccgctg	grgagargrr	gttccgtgct	ctgcgtaaac	atteteactt	geetcaagee	180
139	ctggtcgatg	tegtgggega	cgagagettg	tettataagg	aattttttega	agetaetgte	240
140	ctgttggccc	aatctctgca	taattgcggt	tacaaaatga	acgatgtggt	cagcalligi	300
141	gctgagaata	acacccgctt	tttcatccca	grgarrgeeg	cttggtacat	cygcatgatt	360
142	gtcgcccctg	tgaatgaatc	ttatatccca	gacgagttgt	gcaaggtcat	gggtattage	420
143	aaacctcaaa	tegtgtttac	taccaagaac	attotgaata	aggtettgga	agtgcagtct	
144	cgtactaact	tcatcaagcg	cattatcatt	ctggataccg	tcgagaatat	ccacggctgt	480
145	gagagettge	caaactttat	ttctcgttat	agegaeggta	atategetaa	cttcaagcct	540
146	ctgcattttg	atccagtgga	gcaagtcgcc	gctattttgt	getetagegg	caccaccggt	600
147	ctgcctaaag	gcgtgatgca	gactcaccaa	aatatctgtg	teegettgat	teatgeeetg	660
148	gacccacgtg	tgggtactca	gttgatccct	ggcgtgactg	tcctggtgta	cttgccattc	720
149	tttcacgcct	teggtttttc	tattaccctg	ggctatttca	tggtcggttt	gcgcgtgatc	780
150	atgtttcgtc	gcttcgatca	agaagccttt	ctgaaggcca	ttcaagacta	cgaggtccgt	840
151	agcgtgatca	acgtcccttc	tgtgattttg	ttcctgagca	aatctccatt	ggtcgataag	900
152	tatgacctga	gcagcttgcg	cgaactgtgc	tgtggcgctg	cccctttggc	taaagaggtg	960
153	gccgaagtcg	ctgccaagcg	tctgaatttg	ccaggtatcc	gctgcggctt	tggtctgact	1020
154	gagagcacct	ctgctaacat	tcatagcttg	cgtgatgagt	tcaaatctgg	cagcctgggt	1080
155	cqcqtqactc	ctttgatggc	cgctaagatc	gccgaccgtg	agaccggcaa	agctctgggt	1140
156	ccaaatcaag	tcggcgaatt	gtgtattaag	ggtcctatgg	tgtctaaagg	ctacgtcaac	1200
157	aatgtggagg	ccactaagga	agctattgat	gacgatggtt	ggctgcacag	cggcgacttt	1260
158	ggttattacg	atgaggacga	acatttctat	gtcgtcgatc	gctacaaaga	gttgattaag	1320
159	tataaaggct	ctcaagtcgc	cccagctgag	ctggaagaaa	tcttgctgaa	gaacccttgc	1380
160	attegtgacg	tggccgtcgt	gggtatccca	gatttggaag	ctggcgagct	gcctagcgcc	1440
161	tttgtcgtga	aacaaccagg	caaggaaatt	accgctaaag	aggtctacga	ctatttggcc	1500



Input Set : A:\341-005US1.txt.txt
Output Set: N:\CRF3\09062000\I645706.raw

162	gagogogtgt otcacactaa gta	acctgcgt	ggcggtgtcc	gcttcgtcga	tagcatccct	1560			
163	cgcaatgtca ccggcaaaat tac	ctogtaag	gagttgctga	aacagttgct	ggaaaaggct	1620			
164	qqtqqc					1626			
	<210> SEQ ID NO: 5								
167	<211> LENGTH: 1626								
168	<212> TYPE: DNA								
	<213> ORGANISM: Artificia	al Sequen	ice						
	<pre></pre> <pre><</pre>								
172	2 <223> OTHER INFORMATION: Sequence of a synthetic luciferase								
	4 <400> SEQUENCE: 5								
175	atgatgaaac gcgaaaagaa cgt	gatetae	ggcccagaac	cactgcatcc	actggaagac	60			
176	ctcaccgctq qtqaqatqct qtt	ccatacc	ctqcqtaaac	atagccacct	gcctcaagct	120			
	ctcqtqqacq tcqtqqqtqa cqa					180			
	ctgttggccc aaagcctgca taa		_	_		240			
	gctgagaata acactcgctt ttt					300			
	gtcgccctg tgaatgaatc tta					360			
	aaacctcaaa tcgtctttac tac					420			
	cgtactaatt tcatcaaacg cat		-			480			
	gagagettge ctaactttat cto					540			
	ctgcattttg atccagtcga gca					600			
	ttqcctaaaq qtqtcatqca qac					660			
	gaccetegtg tgggtactea att					720			
	tttcacgcct ttggtttttc tat					780			
	atqtttcqtc gcttcqacca aqa					840			
	totgtgatca atgtoccate tgt					900			
	tatgatetga geagettgeg tga					960			
	geogaggteg etgetaageg tet					1020			
	gagageactt ctgccaacat cca					1080			
	egegtgacee etttgatgge tge					1140			
	ccaaatcagg tcggtgaatt gtg					1200			
	aatgtggagg ccactaaaga ago					1260			
	ggttattatg atgaggacga aca					1320			
	tacaaagget etcaagtege coo					1380			
						1440			
	attegegacg tggccgtcgt ggg					1500			
	tttgtggtga aacaacctgg caa					1560			
	gagegegtgt etcacactaa ata					1620			
	cgcaacgtca ccggcaagat cac	ccycaaa	gaguegeega	aacaaccycc	cyaaaaaycı	1626			
	ggcggc					1020			
	<210> SEQ ID NO: 6								
	<211> LENGTH: 1626								
	<212> TYPE: DNA	1 Caauan							
	7 <213> ORGANISM: Artificial Sequence								
	09 <220> FEATURE:								
	10 <223> OTHER INFORMATION: Sequence of a synthetic luciferase								
	<400> SEQUENCE: 6		~~~~~~~	asataastee	20+0022000	60			
	atgatgaaac gcgaaaagaa cgt					120			
	ctcaccgctg gtgagatgct ctt								
	ctcgtggacg tcgtgggaga cga		-	_		180			
∠10	ctgttggccc aaagcctcca taa	ictgtgga	cacaaaatga	acgatgtggt	gageatttgt	240			

PATENT APPLICATION: US/09/645,706 TIME: 12:02:11 Input Set : A:\341-005US1.txt.txt Output Set: N:\CRF3\09062000\1645706.raw 300 217 gctgagaata acactcgctt ctttatccct gttatcgctg cttggtacat cggcatgatt 360 218 gtcgcccctg tgaatgaatc ttacatccca gatgagctgt gtaaggttat gggtattagc 420 219 aaacctcaaa tcgtctttac taccaaaaat atcctgaata aggtcttgga agtccagtct 220 cgtactaact tcatcaaacg catcattatt ctggataccg tcgaaaacat ccatggctgt 480 221 gagageetge ctaactteat etetegttae agegatggta atategetaa ttteaaacea 540 600 222 etgeattttg atceagtega geaagtggee getattttgt getetteegg eaceaetggt 223 ttgcctaaag gtgtcatgca gactcaccag aatatctgtg tgcgtttgat ccacgctctc 660 224 gaccetegtg tgggtactca attgatecet ggegtgaetg tgetggtgta tetgeettte 720 225 tttcacgcct ttggtttttc tattaccctg ggctatttca tggtcggctt gcgtgtcatc 780 226 atgtttegte gettegaeca agaageette ttgaaggeta tteaagaeta egaggtgegt 840 227 tetgteatea atgteeette agteattitg tieetgagea aateteetti ggitgaeaag 900 228 tatgatotga goagottgog tgagetgtgo tgtggogotg etectttggo caaagaagtg 960 229 geogaggteg etgetaageg tetgaacete eetggtatee getgeggttt tggtttgaet 1020 230 gagagcactt ctgctaacat ccatagcttg cgagacgagt ttaagtctgg tagcctgggt 1080 231 egeqtqacte etettatqqe tqcaaaqate geegaceqtg agaceggcaa agcactggge 1140 232 ccaaatcaag toggtgaatt gtgtattaag ggccctatgg tototaaagg ctacgtgaac 1200 233 aatgtggagg ccactaaaga agccattgat gatgatggct ggctccatag cggcgacttc 1260 234 ggttactatg atgaggacga acacttctat gtggtcgatc gctacaaaga attgattaag 1320 235 tacaaagget etcaagtege eccageegaa etggaagaaa ttttgetgaa gaaceettgt 1380 236 atccgcgacg tggccgtcgt gggtatccca gacttggaag ctggtgagtt gcctagcgcc 1440 237 tttgtggtga aacaacctgg aaaggagatc actgctaagg aggtctacga ctatttggcc 1500 1560 238 gagegegtgt eteacaceaa atatetgegt ggeggegtee gettegtega ttecateeca 239 cgcaacgtga ccggtaagat cactcgtaaa gaattgctga agcaactcct cgaaaaagct 1620 1626 240 gacaac 242 <210> SEQ ID NO: 7 243 <211> LENGTH: 1626 244 <212> TYPE: DNA 245 <213> ORGANISM: Artificial Sequence 247 <220> FEATURE: 248 <223> OTHER INFORMATION: Sequence of a synthetic luciferase 250 <400> SEQUENCE: 7 251 atgatgaaac gcgaaaagaa cgtgatctac ggcccagaac cactgcatcc actggaagac 252 ctcaccqctq qtqaqatqct cttccqaqca ctqcgtaaac atagtcacct ccctcaagca 120 253 ctcgtggacg tcgtgggaga cgagagcctc tcctacaaag aatttttcga agctactgtg 180 254 ctgttqqccc aaaqcctcca taattqtqqq tacaaaatqa acgatgtggt gagcatttgt 240 255 gctgagaata acactegett etttatteet gtaategetg ettggtacat eggeatgatt 300 256 gtcgcccctg tgaatgaatc ttacatccca gatgagctgt gtaaggttat gggtattagc 257 aaacctcaaa tcqtctttac taccaaaaac atcttgaata aggtcttgga agtccagtct 420 258 cgtactaact tcatcaaacg catcattatt ctggataccg tcgaaaacat ccacggctgt 259 gagageetee etaaetteat etetegttae agegatggta atategetaa titteaageee 540 260 tigcattitg atccagtoga gcaagtggcc gctattitgt gctcctccgg caccactggt 261 ttgcctaaag gtgtcatgca gactcaccag aatatctgtg tgcgtttgat ccacgctctc 262 gaccctcgtg tgggtactca attgatccct ggcgtgactg tgctggtgta tctgcctttc 660 263 tttcacgect ttggtttctc tattaccctg ggctatttca tggtcggctt gcgtgtcatc 264 atgtttcgtc gcttcgacca agaagccttc ttgaaggcta ttcaagacta cgaggtgcgt 780 265 teegtgatea aegteeette agteattttg tteetgagea aateteettt ggttgacaag 900 960 266 tatgatetga geagettgeg tgagetgtge tgtggegetg etectttgge caaagaagtg 1020 267 geogaggteg etgetaageg tetgaacete eetggtatee getgeggttt tggtttgact 268 gagagcactt ctgctaacat ccatagcttg cgagacgagt ttaagtctgg tagcctgggt

RAW SEQUENCE LISTING

ZF.Y.1.

DATE: 09/06/2000

<u>Please Note:</u>

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

. . .

DATE: 09/06/2000 TIME: 12:02:12

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/645,706

Input Set : A:\341-005US1.txt.txt Output Set: N:\CRF3\09062000\1645706.raw

L:13 M:270 C: Current Application Number differs, Replaced Current Application No L:13 M:271 C: Current Filing Date differs, Replaced Current Filing Date L:4423 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:228 L:4581 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:242 L:4618 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:245